植 物 分 类 与 资 源 学 报　2012, **34** (6): 592～606
Plant Diversity and Resources　DOI: 10.3724/SP.J.1143.2012.12138

关于植物 DNA 条形码研究技术规范*

高连明[1], 刘　杰[1], 蔡　杰[2], 杨俊波[2], 张　挺[2], 李德铢[1,2]**
(1 中国科学院昆明植物研究所生物多样性与生物地理学重点实验室, 云南 昆明　650201;
2 中国科学院昆明植物研究所中国西南野生生物种质资源库, 云南 昆明　650201)

摘要: DNA 条形码是利用标准的基因片段对物种进行快速鉴定的技术, 已经成功用于生物物种分类和鉴定、生态学调查和生物多样性评估等研究领域。尽管生命条形码数据 (BOLD) 系统提供了主要针对动物类群 DNA 条形码研究的技术规范, 但由于植物本身的生物学特性与所使用的条形码不同, 因此已有技术规范并不完全适用于植物 DNA 条形码的研究。本文根据植物 DNA 条形码研究的特点与我国的实际情况, 编写了植物 DNA 条形码研究技术标准和规范指南, 具体包括十个方面的内容, 即植物 DNA 条形码研究的样品采集策略; 植物标本和野外数据的采集规范; 植物标本图像信息的采集规范; 植物 DNA 材料的采集规范; 植物 DNA 材料的干燥与保存规范; 植物总 DNA 的质量标准及保存规范; 植物标准 DNA 条形码的选择与通用引物; DNA 条形码的扩增与测序; DNA 条形码数据的命名、编辑和提交规范; 以及 DNA 条形码数据分析。我们期望通过这些标准规范的实施和在实践中的不断修订和完善, 能为我国学者开展植物 DNA 条形码和 iFlora 研究提供参考和借鉴。

关键词: 植物 DNA 条形码; 技术规范; 物种鉴定; 标准; 新一代植物志
中图分类号: Q 94-33, Q 948.2　**文献标识码**: A　**文章编号**: 2095-0845(2012)06-592-15

A Synopsis of Technical Notes on the Standards for Plant DNA Barcoding

GAO Lian-Ming[1], LIU Jie[1], CAI Jie[2], YANG Jun-Bo[2], ZHANG Ting[2], LI De-Zhu[1,2]**
(1 *Key Laboratory of Biodiversity and Biogeography*, *Kunming Institute of Botany*, *Chinese Academy of Sciences*, Kunming 650201, China; 2 *Germplasm Bank of Wild Species*, *Kunming Institute of Botany*, *Chinese Academy of Sciences*, Kunming 650201, China)

Abstract: DNA barcoding is a technique using one or a few standardized DNA regions from different genomes for rapid species identification, which is used in the field of taxonomy, ecological surveys and assessment of biodiversity. Because of the plant natural particularity and the DNA barcodes used for plants differing from animals, the standards provided by BOLD which was initially dssigned for animals are not totally compatible in plants DNA barcoding research. Thus, we synthesize and customize a synopsis of technical notes and standards with reference of the BOLD criteria and experience of plant DNA barcoding projects, especial for the researchers with particular interest in plants DNA barcoding in China. Ten aspects related to plants DNA barcoding are covered: 1) sampling strategy for a plant DNA barcoding study, 2) collecting standards for vouches and associate information, 3) collection standards for specimen-referenced images, 4) collecting standards for DNA material, 5) standards for drying and preserving DNA material, 6) quality control and preservation procedures for extracted total genomic DNA, 7) recommended plant barcodes

* 基金项目: 中国科学院大科学装置开放研究项目 (2009-LSFGBOWS-01); 云南省中青年学术技术带头人后备人才项目 (2008PY064); 国家科技部科技基础工作专项项目; 国家高科技研究发展计划 (863 计划) (2012AA021801)
** 通讯作者: Author for correspondence; E-mail: dzl@mail.kib.ac.cn
收稿日期: 2012-11-05, 2012-11-09 接受发表
作者简介: 高连明 (1972-) 男, 博士、副研究员, 主要从事植物系统发育、群体遗传和 DNA 条形码等研究。
E-mail: gaolm@mail.kib.ac.cn

and universal primers, 8) procedures for PCR amplification and sequencing of the DNA barcodes, 9) naming, editing and submission for DNA barcoding files, and 10) procedures and methods for the analysis of DNA barcoding data.

Key words: Plant DNA barcoding; Technical notes; Species identification; Standards; iFlora

DNA 条形码技术是利用标准的基因片段对物种进行快速鉴定（Hebert 等，2003）。该技术提供了可信息化的分类学标准和有效的分类学手段，已经被成功用于生物物种鉴定和分类（Hebert 等，2004；Liu 等，2011），生物多样性调查（Lahaye 等，2008）和生态学研究（Valentini 等，2009）等领域，并成为进展最迅速的学科前沿之一。随着贸易全球化的推进、全球气候变化的加剧，对生物多样性的探索、认识和理解已成为人类可持续发展的重要物质基础，物种的快速鉴定及对生物多样性资源的可持续利用已成为世界性的重大需求（Che 等，2010）。准确的物种鉴定是人类认知生物多样性和可持续发展的必要前提，而基于传统的形态学特征进行物种鉴定已难以满足科学发展的需求，DNA 条形码技术正是在这个背景下应运而生。

植物 DNA 条形码不仅是传统植物分类与鉴定的强有力补充，而且能使标本鉴定过程实现自动化和标准化，突破了对经验的过度依赖，并能够在较短时间内建成易于利用的应用系统，在物种分类和鉴定方面展示出了强大的生命力（Hebert 等，2004；Li 等，2011b，c）。DNA 条形码最大的优点在于可利用生物残迹来鉴定形态学无法鉴定的已知或未知物种（http://www.barcodeoflife.org），因此，DNA 条形码技术在生物多样性调查与监测、分子系统发育与进化、生态学、食品安全、生物检验检疫、法医学、流行病学等领域具有广阔的应用前景。

iFlora 是整合现代植物学、DNA 测序技术与信息技术的集成装备和信息平台，能准确快速识别植物，并能提供相关数字化信息的智能植物志（李德铢等，2012）。利用 DNA 条形码标准数据库的构建真正实现对物种的快速鉴定，同时通过信息化平台实现相关信息获取。由此可见，DNA 条形码不仅是 iFlora 构建的核心内容，更是联系其它相关信息的纽带。

目前，我国虽然有许多学者从事植物 DNA 条形码及相关研究，但在从事该类研究过程中尚缺乏统一的标准和规范可循，致使很多研究尚不规范，不能满足国际生命条形码计划的相关要求，使得目前获得的大量 DNA 条形码数据不能用作标准的参考条码，在一定程度上造成资源的浪费。此外，虽然 BOLD（The Barcode of Life Data）系统提供了 DNA 条形码研究方面的技术规范与要求（Ratnasingham 和 Hebert，2007），但这些技术规范主要针对动物类群，由于植物与动物的自然特性差异较大，所采用的条形码与相应规范不可能完全一致，这对植物 DNA 条形码的应用具有很大的局限性。因此，我们根据近几年的研究实践，针对植物 DNA 条形码本身的特点与研究实际，编写了植物 DNA 条形码研究技术标准和规范，旨在为我国学者进一步开展植物 DNA 条形码研究和 iFlora 构建提供参考和指导。

1　植物 DNA 条形码研究的样品采集策略

自 DNA 条形码提出后，该技术虽然在生物多样性调查和编目、物种鉴定及新物种发现等方面已得到广泛应用（Li 等，2011a），但是需要多少标本（样品）才可以建立一个可靠的物种鉴定参考数据库（Reference library）用于解决物种鉴定的问题，尚未得到很好解决（Zhang 等，2010）。事实上，样品采集的数量应最大限度地覆盖物种的整个遗传变异范围，这样的物种参考数据库才能真正代表该物种的遗传多样性，利用 DNA 条形码鉴定的物种才更可靠。为了获取更多物种的条形码序列，往往就会牺牲物种的样品数量（Matz 和 Nielsen，2005），例如，目前建立的 DNA 条形码数据库中，大多数的物种只包括 5 ~10 条序列，且少数物种仅有 1 ~ 2 条（http://www.barcodinglife.org/views/ligin.php）。国际生命条形码计划（iBOL Project）要求每个物种要有 10 个标本（样品），但这种要求是否合适并没有得到验证。在动物的 DNA 条形码研究中，Matz 和 Nielsen（2005）建议每个物种样本数为 12 个个体。然而 Zhang 等（2010）的研究表明，通常认为的每个物种 5 ~ 10 个样品是不够的，不

能代表物种的遗传多样性，而需要更多的样品数。Liu等（2012）通过对红豆杉属物种DNA条形码采集策略的研究，提出居群内的遗传变异通常小于居群间的遗传变异，建议在采集时应尽可能采集覆盖物种整个分布区不同居群的样品，而每个居群只需要1个样品；来自整个分布区的8~10个不同居群的样品（个体）能够代表物种的遗传多样性，用于植物DNA条形码研究似乎已经足够了。这可能与物种进化的历史及DNA条形码的进化速率有关（Liu等，2012）。因为植物的DNA条形码进化速率比动物的通用条码*COI*要慢，因此，每个物种需要的样品数也是不同的。最近的研究表明，取样的地理范围对DNA条形码的全球应用具有关键的影响（Bergsten等，2012），因此，在DNA条形码采样时考虑物种的地理分布范围至关重要。

DNA条形码样品采集策略的制定需要根据研究目标来确定，通常不同的研究目标，样品的采集策略也不尽相同。目前关于植物DNA条形码的研究主要有两类，一是对特定科属的DNA条形码研究，即建立特定科属内物种的DNA条形码参考数据库，用于物种的鉴定和新物种的发现等；二是基于生态样方或区域的DNA条形码研究，即开展特定区域内的物种组成及生物多样性调查。

基于特定科属的DNA条形码研究，每个物种采集的样品数量与物种的分布范围相关，物种的分布范围越广，则需要采集的样品数量越多，这样才能最大限度地代表物种的遗传变异范围。对于植物DNA条形码研究，每个物种需要的样品数为8~10个个体，且至少来自5个不同的地域或居群。因为来自不同地域和居群的样品（个体）比来自相同地域和居群的多个样品（个体）更能代表物种的遗传多样性，所以用于DNA条形码研究的材料，最好采集来自不同地域和居群的样品，其来源应尽可能地覆盖物种的整个分布区。最好在分布区的中心或主要区域采集，尽量避免从分布区最边缘的区域采集。对于非狭域分布的物种，每个居群不多于2个个体，且个体间应保持一定的空间距离或来自不同的生境，形态特征上也有一定的代表性。不同分布范围的物种建议采集的样品数量见表1。

植物DNA条形码的研究材料可以通过野外采集和馆藏标本检视等途径获得。选择研究样品时，应尽量选择能代表该物种典型形态特征的标本，如果能使用模式标本最好（目前大多数标本馆是不允许的），但来自模式标本产地的植物材料也非常重要。通常馆藏植物标本的DNA降解比较严重，较难获得可用于条形码研究的DNA，因此，使用植物馆藏标本有一定的局限性。野外采集是获取DNA条形码研究材料的主要途径之一。用于植物DNA条形码的样品原则上不能使用植物园或公园中栽培的植物材料，但对于直接从野外移栽到植物园或公园且来源清楚的植物材料可谨慎使用。

基于样方或特定区域的DNA条形码研究，其采集的策略与前者不同。因为特定区域或样方内的物种，其种内遗传变异不大，每个物种采集2~3个个体即可以代表该物种的遗传变异范围。因此，我们建议每个物种取2个个体，但不同的个体间最好保持一定的空间距离（如不同的海拔）或来源于不同的小生境。

2　植物标本和野外数据的采集规范

物种的准确鉴定是建立一个可靠的物种鉴定参考数据库的基础和核心，因此采集完整、具有重要鉴别特征的凭证标本对于开展DNA条形码研究至关重要，这是物种准确鉴定的基础和保障。野外采集植物DNA条形码材料应包括物种的凭证标本、提取基因组DNA的植物材料和植物本身特征及生境的图像或图片信息等。每份样

表1　不同分布范围的物种用于DNA条形码研究的样品采集策略

Table 1　Sampling strategy for plant DNA barcoding according to distribution range

分布范围	样品数量（个体）	取样来源	举例说明
广布种	10~12	不同区域或居群	全球或洲际分布
区域分布	10	不同区域或居群	东亚或喜马拉雅分布
地区分布	8~10	不同区域或居群	华东、华南等
狭域分布	5	不同亚居群	玉龙雪山或金平县

品（或每号标本）采集 3 份标本，用于国内外主要标本馆间的交换与鉴定。采集的凭证标本必须是具有物种重要鉴别特征的完整标本，应包括根、茎、叶、花、果和其它鉴别特征（如芽、鳞茎、球根等），完整的标本有助于对标本的准确鉴定。由于植物与动物的特性不同，标本压制后，很多鉴别特征会丢失，如花色和花形等性状。因此，在野外需要对标本的重要特征进行记录和拍照，并进行初步的鉴定（至少鉴定到属）。最终，再请相关的分类学家进行物种鉴定。对于体型较大的植物（如乔木和灌木等），要确保采集的若干份凭证标本来自同一植株；对于体型较小的植物（如星叶草 *Circaeaster agristis*），最好一份标本采集多个植株，但要保证这些不同的植株均来自同一个小居群。如何采集凭证标本可以参考《标本馆手册》（Bridson 和 Forma，1992）。

采集的标本还需要填写相关的野外采集信息，如采集地点、生境、压制标本后容易丢失的特征（如花的颜色、气味等）和鉴别特征等。BOLD 提供了一个标准的“标本信息电子表格”（Ratnasingham 和 Hebert，2007）。由于 BOLD 提供的“标本信息电子表格”主要依据动物的特性设计的，有些信息并不适用于植物标本。因此，我们根据实际工作需要和植物本身的生物学特性，在 BOLD 提供的“标本信息电子表格”的基础上进行了一些必要的调整和补充。标本采集信息电子表格文件由三个部分内容组成（附表 1，http://journal.kib.ac.cn/UserFiles/File/Excel%20of%20field%20collection%20information(1).xls），各部分所要求的信息分述如下。

2.1　标本采集信息（Collection Data）（表 2）

采集编号（Collection Number）：采集编号是指样品在野外采集过程中由采集者对每份样品给予的编号。每份样品具有唯一的采集编号，我们建议使用采集人的单位英文缩写+采集负责人拼音首字母+数字（流水号）的形式，如中国科学院昆明植物研究所（KIB）高连明（GLM）采集的标本可编号为：KIBGLM0001，中国科学院华南植物园（SCBG）葛学军（GXJ）采集的标本可编号为：SCBGGXJ0002，最好不要采用仅包括数字的采集编号，如 001，232 等，这样的编号很容易与其他人的采集编号相同，而造成后期数据管理的不便和混淆。

采集人（Collectors）：是指实际采集标本的人员。英文书写方式按姓在前名在后的方式，如高连明，Gao Lian-Ming。

采集时间（Collection date）：书写方式为“日-月-年”，月用三个字母缩写，如采集时间为 2012 年 9 月 2 日，写成 02-Sep-2012。

采集地信息（Locality information）：

包括以下字段信息

国家（Country）：采集地所属的国家

省（Province）：填省、自治区和直辖市一级，按国家区划的省级拼音写法书写，如云南省（Yunnan），西藏自治区（Xizang）。

地区（Prefecture/Region）：行政区一级，按国家区划的行政区拼音写法书写，如昆明（Kunming），红河（Honghe）等。

区县（Country/District）：县区（县级市）一级，如石林县、五华区、安宁市等。

具体地点（Exact Site）：具体的采集地点描述，至少到乡镇级或山头，如双龙乡小河村后山，玉龙雪山第一峰。

纬度（Latitude）：以十进制，小数点后保留

表 2　标本采集信息表

Table 2　Excel file of collection data information

附表1：Excel of field collection information.xls [兼容模式] - Microsoft Excel

标本采集信息

采集编号	采集者	采集日期	国家	省自治区	地区	区县	具体地点	纬度	经度	海拔(米)	生境	植物习性	植株高度	其它描述	备注	野外鉴定	DNA材料是否采集
KIBGLM2649	高连明	16-May-2010	中国	云南	德钦	香格里拉	县城至格咱途中20KM	27°59'17.2"N	99°42'15.8"E	3430	高山杂木林	灌木	1-2m	花紫红色，花序5-7花，外疏被鳞片	3	红棕杜鹃	DNA Sampled
KIBGLM2652	高连明	16-May-2010	中国	云南	德钦	香格里拉	县城至格咱途中25KM	28°01'10.4"N	99°43'01.3"E	3330	高山草甸灌丛	灌木	0.5-1m	花兰紫色，花序5-7花，雄蕊10，花萼小	2	灰背杜鹃	DNA Sampled
KIBGLM2653	高连明	16-May-2010	中国	云南	德钦	香格里拉	县城至格咱途中28KM	28°02'26.7"N	99°44'24.4"E	3186	云南松林下	灌木	0.3-1m	花红色或粉红色，花序2-3花生叶腋，雄蕊6-8枚	2	腋花杜鹃	DNA Sampled
KIBGLM2654	高连明	16-May-2010	中国	云南	德钦	香格里拉	格咱至小雪山途中	28°16'07.6"N	99°45'43.1"E	3116	云南松林下	灌木	1.5-2m	花紫红色，花序3花，内有紫红色斑点	2	云南杜鹃	DNA Sampled
KIBGLM2655	高连明	16-May-2010	中国	云南	德钦	香格里拉	小雪山哑口附近	28°18'15.0"N	99°45'32.8"E	3580	针叶林下	灌木	2-3m	花紫红色，花序5-8花	2	红棕杜鹃	DNA Sampled
KIBGLM2656	高连明	16-May-2010	中国	云南	德钦	香格里拉	小雪山哑口	28°19'18.9"N	99°45'36.9"E	3770	杜鹃灌丛	灌木	1-2m	花粉红色，花序20朵左右	2	宽钟杜鹃	DNA Sampled
KIBGLM2657	高连明	16-May-2010	中国	四川	甘孜	乡城	大雪山哑口附近	28°36'12.4"N	99°49'58.9"E	3970	针叶林与杜鹃林	灌木	2-3m	花白色，花序10花，内有紫红色点子，雄蕊10，	2	川滇杜鹃	DNA Sampled

采集信息　采集标签　鉴定标签

4位，如29.4300；或按度-分-秒的形式，如29°25′48″ N。

经度（Longitude）：以十进制，小数点后保留4位，如102.3956；或按度-分-秒的形式，如102°23′44″ E。

海拔（Altitude）：单位米（m），如3 430，3 950。

生境（Habitat）：所采集样品的生长环境，如常绿阔叶林内，高山草甸等。

标本信息（Specimen details）

植物习性（Plant Habit）：乔木（tree），灌木（shrub），半灌木（suffrutex），木质藤本（liana），草质藤本（vine），一年生草本（annual herb），两年生草本（biennial herb），多年生草本（perennial herb），附生植物（epiphyte），寄生植物（parasitic plant）。

植株高度（Height）：所采集植株的高度，使用标准的长度单位，如米（m），厘米（cm）等。

其它描述（Description）：用于描述标本突出的鉴别性状或其它重要信息（如雌雄同株、雌雄异株等）。

备注（Notes）：可以用于记录标本的状况，如一般标本，正模标本或副模标本等信息，也可以记录标本采集的份数等信息。

野外鉴定（Field identification）：在野外鉴定的物种暂定名。

DNA材料是否采集（DNA sample）：注明是否采集了DNA材料。

整理好的采集信息需要按照一定的格式排版后，制作成采集标签，与标本一起进行装订。采集标签的格式可参考图1。

中国西南野生生物种质资源库
Germplasm Bank of Wild Species

采集编号 (COLL.NO.)：KIBGLM2654
采集日期 (DATE)：16-May-2010
采集者 (COLLECTOR)：高连明；马朋飞；蔡兆明
采集地 (LOCALITY)：
中国云南德钦香格里拉格咱至小雪山途中
经纬度：99°45′43.1″ E, 28°16′07.6″ N
海拔 (ALT.)：3 116 m
生境 (HABITAT)：云南松林下
其他描述 (DESCRIPTION)：
灌木，1.5~2 m。花紫红色，花序3花，内有紫红色斑点，雄蕊10枚，伸出花冠外。

DNA Sampled
野外鉴定 (FIELD I. D.)：云南杜鹃

图1　标本的采集标签

Fig. 1　Example of the collecting label

2.2　标本鉴定信息（Specimen identification information）（如表3）

分类学信息包括样品所在的目、科、属和种。按拉丁语正规写法书写。种（Species）的书写按双名法的格式书写，即“属名+种加词”及命名人信息，如 *Rhododendron decorum* Franch.。此外，为便于中国学者之间的交流，还增加了科中文名和种中文名的字段信息（表4）。

物种的名称：虽然用于植物物种鉴定的工具书很多，但为了避免不同的工具书之间存在的异名问题，种子植物的物种鉴定名统一以“Flora of China”为准，苔藓植物的物种鉴定名称以“Bryophyflora of China”为准，其他孢子植物以现有《中国孢子植物志》为准。“Flora of China”出版后如有新的增补或修订的类群，可将最新的专著和修订作为重要参考。在科及以上等级的归属，裸子植物按 Christenhusz 等（2011）的系统，被子植物按 APG Ⅲ系统（APG，2009）。

表3　标本鉴定信息表

Table 3　Excel file of specimen identification information

附表1：Excel of field collection information.xls [兼容模式]

	S	T	U	V	W	X	Y	Z	AA	AB
1	标本鉴定信息									
2	目	科中文名	科	属	种中文名	种	鉴定日期	鉴定者	鉴定者单位	鉴定者联系方式
3	Ericales	杜鹃花科	Ericaceae	Rhododendron	红棕杜鹃	Rhododendron rubiginosum Franch.	2011-05-20	高连明（L.M. Gao）	中国科学院昆明植物研究所	gaolm@mail.kib.ac.cn
4	Ericales	杜鹃花科	Ericaceae	Rhododendron	灰背杜鹃	Rhododendron hippophaeoides Balf. f.et W. W. Smith	2011-05-20	高连明（L.M. Gao）	中国科学院昆明植物研究所	gaolm@mail.kib.ac.cn
5	Ericales	杜鹃花科	Ericaceae	Rhododendron	腋花杜鹃	Rhododendron racemosum Franch.	2011-05-20	高连明（L.M. Gao）	中国科学院昆明植物研究所	gaolm@mail.kib.ac.cn
6	Ericales	杜鹃花科	Ericaceae	Rhododendron	云南杜鹃	Rhododendron yunnanense Franch.	2011-05-20	高连明（L.M. Gao）	中国科学院昆明植物研究所	gaolm@mail.kib.ac.cn
7	Ericales	杜鹃花科	Ericaceae	Rhododendron	红棕杜鹃	Rhododendron rubiginosum Franch.	2011-05-20	高连明（L.M. Gao）	中国科学院昆明植物研究所	gaolm@mail.kib.ac.cn
8	Ericales	杜鹃花科	Ericaceae	Rhododendron	宽钟杜鹃	Rhododendron beesianum Diels	2011-05-20	高连明（L.M. Gao）	中国科学院昆明植物研究所	gaolm@mail.kib.ac.cn
9	Ericales	杜鹃花科	Ericaceae	Rhododendron	川滇杜鹃	Rhododendron traillianum Forrest et W. W. Smith.	2011-05-20	高连明（L.M. Gao）	中国科学院昆明植物研究所	gaolm@mail.kib.ac.cn
10										

样品鉴定人为该样品的具体鉴定人，书写方式如前所述；鉴定日期是指鉴定该标本的具体日期，格式如标本采集日期；联系方式和所在单位原则上是样品鉴定人的联系方式（电子邮件地址）与所在单位全称；如果使用的样品是标本馆的馆藏标本，应提供可以联系到的具体人员的联系方式，以便检视到该标本。

凭证标本经过专家鉴定后，需要附上鉴定标签。鉴定标签中包括的内容除物种的中文名、拉丁名、鉴定人、鉴定日期等外，还应包括采集编号，以免粘贴鉴定标签时“张冠李戴”。鉴定标签的格式可参考图 2。

采集编号 (Coll.No.)：KIBGLM2654
杜鹃花科 Ericaceae
云南杜鹃
Rhododendron yunnanense **Franch.**
鉴定人 (Det.)：高连明 (L.M. Gao) 2011-5-20

图 2　标本的鉴定标签

Fig. 2　Example of the identification label

2.3　标本凭证信息（Voucher information）

（附表 1，参看以下链接：http://journal.kib.ac.cn/UserFiles/File/Excel%20of%20field%20collection%20information(1).xls）

样品编号（Sample ID）：样品编号是样品在进行 DNA 条形码实验过程中用到的唯一编号，可由字母和数字组成。作者可以根据自己所做的类群和项目的名称代码来编号，但应该便于记忆且最好为连续的编号，强烈建议使用与采集编号相同的编号。

凭证标本编号（Museum ID）：凭证标本编号是按样品所存放标本馆的编号系统对应的编号，不同的标本馆有不同的凭证编号系统，在标本馆制作馆藏标本时给予凭证编号。

收藏代码（Collection code）：如为个人收藏，则注明个人收藏代码。

标本存放处（Institution storing）：标本目前存放的标本馆名称或研究机构名称。

样品提供者（Sample donor）：样品的提供者指提供样品的人，如果是研究者，应提供具体的姓名。英文书写方式如前所述。

提供者联系方式（Donor email）：标本提供者的电子邮箱，按标准书写。

以上字段信息中，样品编号和标本存放处为必须填写的信息。

3　植物标本图像信息的采集规范

植物标本图像信息包括植物在野外生长环境中拍摄的图像和图片，能反映该物种的生长环境、伴生物种、植株本身的形态鉴定特征（如叶、花、果及种子等）等信息；也包括凭证标本的图片信息，使用的相机像素不低于 500 万像素，图像清晰，JPG 格式。图片的一般要求：生境照片 1 张、全株特写 1～2 张、形态鉴别特征 3～5 张；图片的像素推荐不小于 2048×1536 或 1536×2048，图像的长宽比例为 3∶2；图像的命名按照采集编号+物种拉丁名+照片内容（如：“生境”、“植株”、“花”、“果”），如果照片内容相同，则在照片内容后加数字加以区分（如 KIBGLM0001-Taxus wallichiana-seed 1，KIBGLM0001-Taxus wallichiana-seed 2），并将同一物种的图片放在一个文件夹内，文件夹使用采集编号命名。

4　植物 DNA 材料的采集规范

植物 DNA 材料指获取植物基因组 DNA 的组织或器官，通常采集植物的叶片材料，当叶片材料难以获取或被真菌或细菌感染时，也可以采集植物的其它组织或器官（如芽、花、果实和种子等）用于获取基因组 DNA。对于叶片容易获得且不易被真菌等感染的植物，采集当年生的近成熟的、健康的新鲜叶片作为提取基因组 DNA 的材料；对于植物叶片不易获得或容易被真菌等感染的植物，可采集植物的花、芽或种子等作为 DNA 材料。每份样品（每号标本）采集 1 份 DNA 材料，DNA 材料应采集 3～5 g 叶片（鲜重）或其它组织和器官材料，或叶表面积不小于 50 cm^2（尺寸大小约等于成年人的手掌面积），每号 DNA 材料最好分为 2 份保存。可以将新鲜的植物 DNA 材料置于柔软且透气性较好的纸袋中进行干燥。透气的纸袋可以将植物 DNA 材料在干燥时与硅胶（干燥剂）分开，这样即便植物叶片因干燥变脆后，也不会在运输过程中被硅胶挤碎，不会与硅胶混杂，并造成样品间的交叉污染。如果没有类似的纸袋也可用塑料的自封袋替

代，但植物DNA材料就需要与硅胶置于一起进行干燥。存放植物DNA材料的袋子上要注明采集编号和采集日期等信息。

由于植物习性多样（如草本、藤本、灌木、小乔木、大乔木等），植物叶片类型复杂（如纸质的、革质的、蜡质的、多浆的、刺状的等），因此很难制定统一的采集规范。但在采集植物DNA材料时应注意以下事项：

4.1 对于体型较大的植物（如乔木、灌木和大型草本植物等），要确保DNA材料与采集的凭证标本来自相同的植株，DNA材料的编号与凭证标本的采集编号必须相同；对于体型较小的植物（小草本或小灌木等），可采集整个植株作为DNA材料，也可采集多个植株作为DNA材料，但不同的植株应采自同一小居群，且不同植株要分开保存和干燥，并赋予不同的编号（如KIBGLM001A、KIBGLM001B、KIBGLM001C），而不能将不同的植株混在一起进行干燥和保存。

4.2 对于易失水的植物DNA材料（如纸质叶或幼嫩叶片等），应在野外采集DNA材料并立即进行快速干燥；对于不易失水的植物遗传物质材料（如革质叶或蜡质叶等），可在采集后放置一段时间（在叶片失水前），然后再进行干燥，但最好在采集后立刻干燥；对于难于用硅胶快速干燥的遗传物质材料（如较厚或具多浆的叶片或种子等），可将这些材料碎成小片（块）后用硅胶快速干燥。

4.3 为保证遗传物质材料能快速干燥，同一份遗传物质材料不易过多。如需采集较多量的遗传物质材料，可分别放入不同的纸袋中进行干燥。

5 植物DNA材料的干燥与保存规范

为确保植物材料中的基因组DNA的完整性，植物DNA材料采集后需要快速干燥，目前快速干燥植物DNA材料的最佳方式是硅胶（一种干燥剂），能够使植物材料在短时间内快速脱水，并且对DNA的损伤较小。干燥用的硅胶最好是粉末状的白色硅胶（直径为20～30 Å）与变色硅胶（蓝色）混合使用，粉末状硅胶能最大限度地与植物材料接触，可快速使植物材料脱水干燥，变色硅胶不但可以脱水，还可以清楚地指示硅胶的干燥状态，了解植物DNA材料的干燥程度，但由于变色硅胶为直径较大的圆形颗粒（通常直径为1.5～3 mm），与植物DNA材料接触面较粉末状的硅胶小，植物材料脱水干燥的时间要长一些。没有粉末状硅胶的情况下也可以完全使用颗粒状的变色硅胶干燥植物DNA材料。

5.1 植物DNA材料的干燥规范

将新采集的植物DNA材料放入透气良好的纸袋中并密封，然后置于密封性能好的干燥箱（或盒），并放入足够的硅胶（使硅胶覆盖全部的遗传物质材料）进行快速干燥；在没有干燥箱（盒）的条件下，将植物遗传物质材料放入透气良好的纸袋中并密封，然后放入塑料的自封袋中并加入足够量的硅胶进行快速干燥；如果使用塑料的密封袋干燥植物DNA材料，则需要将颗粒状的变色硅胶与植物DNA材料放在密封袋一起干燥。如果变色硅胶变为粉红或红色，要及时更换硅胶，直到变色硅胶保持为蓝色，表明遗传物质材料已完全干燥。

5.2 硅胶干燥的植物DNA材料保存规范

干燥好的DNA材料可根据采集编号或保存编号进行排列，将有植物材料的纸袋密封好放入到密封盒中，有开口的一侧朝上，然后再放入少量的变色硅胶进行保存；对于易碎的纸质叶或质地较薄的叶片材料用透气性较好的纸袋保存，并防止相互挤压；定期检查变色硅胶的颜色，保持变色硅胶始终为蓝色，以防止遗传物质材料再度吸水；植物DNA材料保存的条件为：温度（4～15 ℃）、湿度（相对湿度小于10%～30%），如果没有相应的保存条件，可在室温下保存，但一定保持在通风条件良好且干燥的环境下；植物DNA材料也可以在低温（–20 ℃）或超低温（–80 ℃）条件下保存，但需要保持DNA材料的干燥状态；由于植物DNA材料（如叶片等）很难长期保存，如在室温下保存超过2年的植物DNA材料其基因组DNA就会发生降解，因此建议最好保存总DNA。将保存在干燥密封盒中的植物DNA材料编号等信息的标签贴在密封盒外面的相应区域，便于查找或检索。

6 植物总DNA的质量标准及保存规范

目前，植物总DNA提取的方法很多，传统上主要有CTAB法、SDS法、低盐高渗法。近年

来出现了以螯合树脂、特异性 DNA 吸附膜、离子交换纯化柱及磁珠或玻璃粉吸附等为基础的植物 DNA 提取新方法，并开发出了许多种商业用植物总 DNA 提取的试剂盒。研究表明，在植物总 DNA 的提取方法中，CTAB 法和试剂盒法对大部分野生植物均能得到较高质量的 DNA，但有时 CTAB 法得到的总 DNA 纯度较低，还需要进一步纯化，但成本比较便宜；而试剂盒法提取 DNA 操作简单、高效、DNA 质量较高，但价格昂贵，且得率较低（刘杰和高连明，2011）。高速离心密度梯度法是获取大量高质量基因组 DNA 的一种有效的方法，但由于成本昂贵，需要专用的设备，当前使用并不普遍，但是该方法是建立 DNA 库所选用方法的最佳选择。经典的 CTAB 法是由 Doyle 和 Doyle（1987）提出，后经过多位学者的改良，是目前使用最为普遍的一种植物总 DNA 提取方法。

植物总 DNA 作为一种重要的遗传物质材料用于保存，不但要求总 DNA 有很高质量，而且还能够长时间保存，因此，总 DNA 本身在完整性、纯度、浓度和含量等方面需要达到一定的标准和要求。完整性：基因组 DNA 长度大于 15 kb；纯度：$1.7 < A_{260}/A_{280} \leq 2.0$，同时凝胶电泳检测为单一明亮的 DNA 条带，DNA 限制性酶切反应产物琼脂糖凝胶电泳检测为均匀弥散状 DNA；浓度：总 DNA 浓度不小于 100 $ng \cdot \mu L^{-1}$，体积不少于 100 μL；含量：每个保存样本的总 DNA 含量应大于 10 μg。

植物总 DNA 的保存最好以干粉状态保存，如果总 DNA 以溶液形式保存，则用 1×TE 缓冲液稀释，且 TE 的缓冲液的 pH 值为 8.0~8.5 之间；植物总 DNA 低温保存最好在-80 ℃以下或液氮保存，无条件的也可在-20 ℃保存；总 DNA 应避免经常冻融；同时建议每份样品保存 3 ~5 个样本。此外，植物总 DNA 还需要定期检测与更新，每隔 2 年一次，随机抽样检测保存的样品，若检测结果不符合上述总 DNA 的保存要求，则按以上的规范，重新采集、提取、测定与保存样品。

7　植物标准 DNA 条形码的选择与通用引物

虽然 *COI* 基因作为动物的通用条形码已得到了广泛应用，但并不适合作为植物的 DNA 条形码。国际生命条形码联盟植物工作组（CBOL Plant Working Group，2009）通过对 7 个叶绿体 DNA 片段的综合分析评估，最后推荐 *rbc*L 和 *mat*K 组合作为陆地植物的核心 DNA 条形码，用于植物物种鉴定的统一框图。此外，在第三届国际生命条形码大会上，提出 ITS（包括 ITS2）和 *trn*H-*psb*A 可作为植物条形码的辅助条形码，并建议在 18 个月内进行进一步的评估（Hollingsworth 等，2011）。中国植物条形码研究组（China Plant BOL Group）根据对主要来自中国的种子植物 75 科 141 属 1 757 种共约 6 286 个样本的 4 个 DNA 候选条形码片段（*rbc*L，*mat*K，*trn*H-*psb*A 和 ITS）的综合分析，建议 ITS（或 ITS2）也作为种子植物的核心条形码之一，提出 *rbc*L + *mat*K + ITS 组合（Li 等，2011a）。鉴于 *trn*H-*psb*A 片段在多数植物类群具有较高的通用性和物种分辨率，在此我们也提供了该条码的相关信息，便于相关研究人员参考使用。

我们将目前推荐的 4 个植物 DNA 条形码和推荐使用的主要引物的名称、引物序列、来源及适用的类群进行了总结（表 4）。在植物的 4 个 DNA 条形码中，*rbc*L 片段所推荐的 2 对引物均具有很高的通用性，可以选择其中的任何 1 对引物用于条形码的扩增和测序。针对 *mat*K 引物通用性较差的问题，由国际生命条形码植物工作组主席 Hollingsworth 教授领导的 *mat*K 指导组（*mat*K Steering Group）对陆地植物中不同植物类群（如苔藓植物、蕨类植物、裸子植物和被植物等）开发了专门的 *mat*K 引物，并给出了相应的扩增及测序的实验程序和反应条件，相关引物信息见表 4，具体的反应条件和实验程序可参考下面网址中的相关 PDF 文件（http://connect.barcodeoflife.net/group/plants）。ITS 在种子植物中的通用性较低（Li 等，2011a），今后还需要开发在陆地植物中通用性更高的 ITS 引物。表 4 中列出的 ITS 引物是目前种子植物中通用性较高的引物组合，对某些特定类群使用的引物本文中没有列出，研究人员可以参考相关的文献。ITS 片段的三个正向引物（ITS5，ITS5a 和 ITS1）均可以与反向引物 ITS4 组合进行 PCR 反应（表 4），其通用性高低在不同类群间可能存在差异，研究者可根据特定研究类群确定使用哪个引物组合。本文中也提供了一对裸子植物的 ITS 引物组合

表4 四个陆地植物 DNA 条形码的引物信息及来源

Table 4 Details of primers used for the four DNA barcoding regions of land plants

Region	Primer Name	Direction	Sequence 5′-3′	Source	Notes
*rbc*L	rbcLa_f	f	ATGTCACCACAAACAGAGACTAAAGC	Kress & Erickson, 2007	Land plants
	rbcLa_rev	r	GTAAAATCAAGTCCACCRCG	Kress & Erickson, 2007	Land plants
	rbcL_ajf634R	r	GAAACGGTCTCTCCAACGCAT	Fazekas *et al.*, 2008	Land plants
	1F	f	ATGTCACCACAAACAGAAAC	Fay *et al.*, 1997	Land plants
	724R	r	TCGCATGTACCTGCAGTAGC	Fay *et al.*, 1997	Land plants
*mat*K	390F	f	CGATCTATTCATTCAATATTTC	Cuénoud *et al.*, 2002	Angiosperms
	1326R	r	TCTAGCACACGAAAGTCGAAGT	Cuénoud *et al.*, 2002	Angiosperms
	3F_KIM	f	CGTACAGTACTTTTGTGTTTACGAG	Kim unpublished	Angiosperms
	IR_KIM	r	ACCCAGTCCATCTGGAAATCTTGGTTC	Kim unpublished	Angiosperms
	xF	f	TAATTTACGATCAATTCATTC	Ford *et al.* (2009)	Angiosperms
	MALPR1	r	ACAAGAAAGTCGAAGTAT	Dunning & Savolainen (2010)	Angiosperms
	472F	f	CCCRTYCATCTGGAAATCTTGGTTC	Yu *et al.*, 2011	Angiosperms
	1248R	r	GCTRTRATAATGAGAAAGATTTCTGC	Yu *et al.*, 2011	Angiosperms
	Gym_R1A	f	TCAYCCGGARATTTTGGTTCG	Li *et al.*, 2011b	Gymnosperms
	Gym_F1A	r	ATYGYRCTTTTATGTTTACARGC	Li *et al.*, 2011b	Gymnosperms
	Ehp_F	f	TCATTCAGAGCTGTTAGTTAG	Li *et al.*, 2011b	Gnetales
	Ehp_R	r	ATCGTACTTTTATGCTTACAGGC	Li *et al.*, 2011b	Gnetales
	GneF	f	TCATYCRRARCTKTTVATHMG	Royal Botanic Garden Edinburgh	Gnetales
	GneR	r	ATMGTACTTTTATGYTTMMARGC	Royal Botanic Garden Edinburgh	Gnetales
	HORN-F1	f	GCAAGAACGTTTCTTATATCC	Royal Botanic Garden Edinburgh	Hornworts
	HORN-R2	r	TTTRGCACATGAAAATCGAAG	Royal Botanic Garden Edinburgh	Hornworts
	PolypodF1	f	ATTTYTGGARGAYAGAYTDCC	Royal Botanic Garden Edinburgh	Polypodiales
	PolypodF2	f	AATTTCRCARTCYAYYCATTC	Royal Botanic Garden Edinburgh	Polypodiales
	PolypodR1	r	CGTRGTATATATCTCRATYTACGC	Royal Botanic Garden Edinburgh	Polypodiales
	SCL-F	f	GCCGACGAATTCAAGATG	Royal Botanic Garden Edinburgh	Early-diverging Polypodiales
	SCL-R	r	GCTAATTTMSTAASWGGACGTCC	Royal Botanic Garden Edinburgh	Early-diverging Polypodiales
	PTE-F	f	ACTYYAATTCGATTGTTTCG	Royal Botanic Garden Edinburgh	Pteridaceae
	PTE-R	r	AARGAAAAVACTTGCCAAAG	Royal Botanic Garden Edinburgh	Pteridaceae
	TreeFernF	f	CACATMTHTCAAGRTTGGYTCTC	Royal Botanic Garden Edinburgh	Cyatheales
	TreeFernR	r	ATATCTYAATCTACGCAAYCC	Royal Botanic Garden Edinburgh	Cyatheales
	LYC-F	f	CTTATACGAATTTTTCGTCGACG	Royal Botanic Garden Edinburgh	Lycopodiales+Isoetales
	LYC-R	r	TTTTYGCACATGAAAATCG	Royal Botanic Garden Edinburgh	Lycopodiales+Isoetales
	SEL-F	f	TAGRTACCGAATTATCTTACTC	Royal Botanic Garden Edinburgh	Selaginellales
	SEL-R	r	TATGTGTTGCATTCGGTAC	Royal Botanic Garden Edinburgh	Selaginellales
	EQU-F	f	GAATCTTTTATTCGAATTCTTCG	Royal Botanic Garden Edinburgh	Equisetales
	EQU-R	r	GTCGTACTTTTATGTTTACGAGC	Royal Botanic Garden Edinburgh	Equisetales
	OPH-F	f	TMTTATTCGAMTTCYTCGTCG	Royal Botanic Garden Edinburgh	Ophioglossales

续表 4 **Continued table 4**

Region	Primer Name	Direction	Sequence 5′-3′	Source	Notes
*mat*K	OPH-R	r	TTGTACTTTTRTGTTTGCAAGC	Royal Botanic Garden Edinburgh	Ophioglossales
	MAR-F	f	RTTCGAATYTTTCGTCGAC	Royal Botanic Garden Edinburgh	Marattiales+Pasilotales
	MAR-R	r	TCTTATGTTTACAAGCCAACG	Royal Botanic Garden Edinburgh	Marattiales+Pasilotales
	OSM-F	f	GAATCCAGGATGCTTCG	Royal Botanic Garden Edinburgh	Osmundales
	OSM-R	r	ATAATCCATCTCTATCTATTGATCC	Royal Botanic Garden Edinburgh	Osmundales
	HYM-F	f	CRKCRAATCAAGGATGC	Royal Botanic Garden Edinburgh	Hymenophyllales
	HYM-R	r	CATYATCYGTTAACGTAGTCC	Royal Botanic Garden Edinburgh	Hymenophyllales
	GLE-F	f	GCTTYRATTCGAATSTTTCG	Royal Botanic Garden Edinburgh	Gleicheniales
	GLE-R	r	TCTGTHAAAGTAGTCCRAGC	Royal Botanic Garden Edinburgh	Gleicheniales
	SCH-F	f	RAYTCGAAYRTTTCGTCG	Royal Botanic Garden Edinburgh	Schizaeales
	SCH-R	r	TYGTASTTTTATGTTTACAAGC	Royal Botanic Garden Edinburgh	Schizaeales
	SAL-F	f	TTTRATTCGAATGTTTCGTAG	Royal Botanic Garden Edinburgh	Salviniales
	SAL-R	r	KGATTTACWAACAGGATGTCC	Royal Botanic Garden Edinburgh	Salviniales
	matK-LivF1A	f	TYCATCCWGAAATTTTGATTCG	Royal Botanic Garden Edinburgh	Liverworts
	matK-LivR1A	r	ATAGTACTTTTRTGTTTACATGC	Royal Botanic Garden Edinburgh	Liverworts
	LivCTH-442F	f	ATACCTTAYTTTTTTCAYCC	Royal Botanic Garden Edinburgh	Liverworts
	LivCTH-1171R	r	CATTATCTGDTAATGTTGTCC	Royal Botanic Garden Edinburgh	Liverworts
	LivLFY-442F	f	ATACCWTATTCTTTYCAYCC	Royal Botanic Garden Edinburgh	Leafy liverworts
	LivLFY-1171R	r	CATDATCTGYSAAIGTTGTCC	Royal Botanic Garden Edinburgh	Leafy liverworts
	Moss404F	f	GGACTARYTATCAATCTATTCAYTC	Royal Botanic Garden Edinburgh	Moss
	Moss1336R	r	TRCAAGCYAAYGTTTTAGC	Royal Botanic Garden Edinburgh	Moss
	Moss404F	f	GGACTARYTATCAATCTATTCAYTC	Royal Botanic Garden Edinburgh	Moss
	Moss1324R	r	GTTTTAGCACAWGAAAATCG	Royal Botanic Garden Edinburgh	Moss
	Moss485F	f	AAATACCTYATTTTWTTCATCC	Royal Botanic Garden Edinburgh	Moss
	Moss1336R	r	TRCAAGCYAAYGTTTTAGC	Royal Botanic Garden Edinburgh	Moss
ITS	ITS5	f	GGA AGTAAAAGTCGTAACAAGG	White *et al.*, 1990	Land plants
	ITS5a	f	CCTTATCATTTAGAGGAAGGAG	Stanford *et al.*, 2000	Land plants
	ITS1	f	TCCGTAGGTGAACCTGCGG	White *et al.*, 1990	Land plants
	ITS4	r	TCCTCCGCTTATTGATATGC	White *et al.*, 1990	Land plants
	ITS-Leu	f	GTCCACTGAACCTTATCATTTAG	Baum *et al.*, 1998	Gymnosperms
	ITS4	r	TCCTCCGCTTATTGATATGC	White *et al.*, 1990	Gymnosperms
ITS2	S2F	f	ATGCGATACTTGGTGTGAAT	Chiou *et al.*, 2007	Seed plants
	S3R	r	GACGCTTCTCCAGACTACAAT	Chiou *et al.*, 2007	Seed plants
	GYM_5.8S F2	f	GYAGAATCCCGTGARTCATC	This study, designed by GLM*	Gymnosperms
	ITS4	r	TCCTCCGCTTATTGATATGC	White *et al.*, 1990	Gymnosperms
*trn*H-*psb*A	psbAF	f	GTTATGCATGAACGTAATGCTC	Sang *et al.*, 1997	Land plants
	trnH2	r	CGCGCATGGTGGATTCACAATCC	Tate & Simpson 2003	Land plants

* GLM: Lian-Ming Gao is a co-author in this study

(ITS-Leu/ITS4)，其通用性是目前最高的一对引物。裸子植物的 ITS 序列长度变异极大，作为 DNA 条形码并不合适，但其变异区主要存在于 ITS1 基因间区（变异范围为 630～3 125 bp），而 ITS2 的长度变异不大（225～255 bp）(Liston 等，1996；Gernandt 和 Liston，1999)，因此，ITS2 可作为裸子植物的 DNA 条形码（Yao 等，2011）。表 4 中列出的 2 对 ITS2 的引物均具有很高的通用性。辅助条形码 *trn*H-*psb*A 推荐的引物通用性高，可用于绝大部分陆地植物。

8 DNA 条形码的扩增与测序

通过 PCR 扩增和测序过程获取 DNA 条形码序列，建立物种鉴定的参考数据库，以便开展后续的 DNA 条码分析和物种鉴定工作。PCR 扩增成功与否是获得 DNA 条码序列的关键步骤。

8.1 DNA 条形码的扩增

PCR 扩增体系中，聚合酶非常重要。现在使用的聚合酶种类很多，其功能也有所不同，有些可以提高降解 DNA 的 PCR 扩增，如 Restorase 和 Restorase II 酶（Sigma）；有的可以缩短 PCR 扩增反应时间，如 Z *Taq* 酶（Takara)，一次 PCR 扩增反应只需 20 min；有的酶则可以提高 DNA 复制的准确性，如 Puhsion *Taq*（BioLabs)，DeepVent *Taq* 酶（NEB）和 Diamond *Taq* 酶（Bioline）等。因为 Invitrogen 公司生产的 *Taq* 聚合酶（Platinum® *Taq* DNA Polymerase）比其它 *Taq* 酶的扩增效果好，已成为 CCDB 的标准酶，但价格相对较高。我们建议选择可以提高 DNA 复制的准确性的聚合酶，如 Puhsion *Taq*（BioLabs)，不仅可以提高 PCR 的扩增效率，对含有单核苷酸重复序列可以提高序列的读长（Fazekas 等，2010）。国外公司的 Platinum *Taq*（Invitrogen）和 Biolase *Taq*（Bioline)，以及国内天根生化科技有限公司（TIANGEN）的 *Taq* DNA Polymerase 扩增效率和复制准确性也较高。

8.2 PCR 反应体系

在开展 PCR 扩增的过程中，所有的 PCR 试剂都需要使用灭菌消毒的吸头或过虑吸头，以免污染，移取 *Taq* 酶和其它试剂时，要用无菌吸头，实验中还要设置阴性对照和阳性对照。

有关 *mat*K 基因在不同陆地植物类群中的 PCR 扩增与测序的反应体系及反应程序参考 http://connect.barcodeoflife.net/group/plants。本文仅给出其它三个 DNA 条形码的反应体系。我们推荐使用 20 μL 的 PCR 反应体系，具体如表 5。

表 5 PCR 反应体系

Table 5 The reaction system of PCR

Reagents	1 reaction /μL	100 reactions /μL
10×buffer	2.0	200
50 mM $MgCl_2$	1.0	100
10 mM dNTPs	0.8	80
10 μM Forward Primer	0.25	25
10 μM Reverse Primer	0.25	25
Polymerase (5 U/uL)	0.2	20
ddH_2O	4.5	450
Additives (10% Trehalose)	10	1000
Total	19	1900
DNA template ($10-50\ ng \cdot \mu L^{-1}$)	1.0 per reaction	

海藻糖（trehalose）是一种广泛使用的 PCR 优化剂。将它加入到 PCR 反应体系中，即可以降低 DNA 熔融温度，又可以使 *Taq* 聚合酶更稳定，消除 DNA 抑制物对 PCR 扩增的影响，提高 PCR 扩增成功率，还可以增加序列的读取长度。在大多数类群中不使用优化剂也可获得很高的 PCR 扩增效率，但使用优化剂往往会提高 PCR 扩增效率。此外，其它 PCR 优化剂还有 DMSO，BSA 和 Betaine 等，在不同的 DNA 条形码或类群中效果可能不同，可选择使用。

8.3 PCR 反应条件

*rbc*L，*trn*H-*psb*A 和 ITS（ITS2）三个条码的 PCR 反应条件不尽相同，且在不同的类群中所要求的反应条件也可能不同，在此只给出三个 DNA 条码在我们实验室中扩增效率最高的程序，供参考。

***rbc*L 的 PCR 反应程序：**

95 ℃ 4 min；[35 cycles：94 ℃ 30 s；54 ℃ 1 min；72 ℃ 1 min]；72 ℃ 10 min；4 ℃ ∞。

***trn*H-*psb*A 的 PCR 反应程序：**

95 ℃ 4 min；[35 cycles：94 ℃ 30 s；55 ℃ 1 min；72 ℃ 1 min]；72 ℃ 10 min；4 ℃ ∞。

ITS 的 PCR 反应程序：

被子植物：94 ℃ 4 min；[35 cycles：94 ℃ 45 s；

50 ℃ 1 min；72 ℃ 1 min]；72 ℃ 10 min；4 ℃ ∞。

裸子植物：94 ℃ 4 min；[35 cycles：94 ℃ 1 min；52 ℃ 1.5 min；72 ℃ 2 min]；72 ℃ 10 min；4 ℃ ∞。

ITS2 的 PCR 反应程序：

94 ℃ 3 min；[35 cycles：94 ℃ 30 s；52 ℃ 30 s；72 ℃ 45 s]；72 ℃ 7 min；4 ℃ ∞。

8.4 PCR 产物的检测与纯化

PCR 产物可以用琼脂糖凝胶电脉进行检测，大多数实验室都使用这种方法。

为去除 PCR 产物中未结合的核苷酸和剩余引物，需要对 PCR 产物进行纯化。如果不进行产物纯化，那么序列前 50 个左右碱基在测序反应中将无法正确解读，因此，测序反应前必须要对 PCR 产物进行纯化。PCR 产物纯化的方法有多种，如用试剂盒纯化（有很多公司提供相关的产品）、酶消化法纯化等。我们在此推荐使用 ExoSAP-IT（USB Corporation，USA）酶消化法纯化。反应步骤如下：将 2 μL 的 ExoSAP-IT（稀释 10 倍）加入到 5 μL 的 PCR 产物中混合，然后 37 ℃温育 30 min 以降解多余的引物和核苷酸，然后在 80 ℃温育 15 min 使 ExoSAP-IT 酶失活，纯化后的 PCR 产物可直接用于测序反应，这种方法可在 PCR 仪上完成，比使用试剂盒的方法简单快速。

8.5 测序反应体系

为了获取可靠的 DNA 条形码序列，每个 DNA 条形码必须对正反两个引物进行测序。为得到较好的测序结果的同时减少成本，原始的 BigDye 可稀释 24 倍后进行测序反应。如果 PCR 产物浓度过高，可在测序前定量 PCR 产物浓度，加水稀释至 25 ng·μL^{-1}。我们推荐使用 5 μL 的测序反应体系，具体如表 6。

表 6 测序反应体系

Table 6 Sequencing reaction system

Reagents	1 reaction /μL	100 reactions /μL
5×buffer	1.0	100
10 μM primer	0.5	50
BigDye	0.3	30
ddH_2O	2.2	220
Total	4	400
Purified PCR product	1.0 per reaction	

如果在 PCR 扩增反应中使用了优化剂，建议在测序反应中也加入相同的优化剂，加入的比例与 PCR 反应相同，这样通常会提高测序的成功率和序列读长。

8.6 测序反应条件

96 ℃变性 10 s，50 ℃退火 5 s，60 ℃延伸 4 min，共 32 个循环。

8.7 测序产物纯化

目前，测序反应产物纯化方法主要包括酒精沉淀（ethanol precipitation），磁珠吸附（magnetic bead）或交联葡聚糖（Sephadex-based）。磁珠吸附方法适用于高通量设备，具有更高的灵敏度，可以减少测序反应中使用的 BigDye 的用量，从而减少测序成本。而交联葡聚糖方法适用于低或中等通量设备，廉价且可靠。酒精沉淀很多实验室也经常使用，效果也不错。纯化后的测序产物可使用测序仪进行序列测定（如 Applied Biosystems 3730 或 3730xL DNA Analyzer；Applied Biosystems）。

9 条形码数据的命名、编辑和提交规范

因为植物有多个核心 DNA 条形码，为了规范植物 DNA 条形码序列的命名，我们对不同植物 DNA 条形码给予统一的缩写格式（参考 CBOL Plant Working Group，2009；Li 等，2011a）。各条码的缩写为：*rbc*L（R），*mat*K（M），*trn*H-*psb*A（P），ITS（I）和 ITS2（I2）。正反向引物的命名原则为正向引物（Forward）为 F，反向引物（Reverse）为 R。

9.1 DNA 条形码序列的命名

获得 DNA 条形码序列的原始峰图文件（trace file）后，需要对这些原始文件和编辑后的文件进行命名和分析。原始峰图文件的命名原则为：DNA 条形码缩写字母+样品编号（Sample ID）+引物方向，中间以下画线分开。如样品编号为 KIB002 样品的 *rbc*L 片段正反两个引物的原始峰图文件的命名分别为：R_KIB002_F 和 R_KIB002_R。对于编辑好的序列文件的命名原则为：DNA 条形码缩写字母+样品编号（Sample ID），中间以下画线分开，如 R_KIB002。

9.2 序列编辑

序列编辑软件使用最广泛的是 Sequencher™（Gene Codes Corporation），SeqScape®（Applied

Biosystems）和 DNA Star（Lasergene®）。我们建议使用 Sequencher 进行序列编辑，先剪去序列两端不可靠的碱基序列和引物序列，然后再将正反两个引物的序列进行比对（assemble），对序列进行编辑最后获得正反引物序列一致的（consensus）DNA 条形码的序列。编辑好的 DNA 条形码序列保存为 fasta 格式（*.fasta）。

条码序列长度原则上越长越好（紧接着引物开始），最好不要短于引物后的 20 bp。如果反向引物能够测通整条序列，建议 DNA 条形码序列从引物区以后保留，如果反向引物不能够测通整条序列，建议在引物区后尽量保留更多的可靠碱基的序列，即在不确定区之后开始保留，如图 3。

9.3 序列比对

目前很多序列比对软件，如 Clustal，Bioedit 和 Mega 等都可用于条形码基因序列比对，可根据自己的喜好选择使用。

9.4 序列质量与矩阵核查

获得的 DNA 条形码原始峰图应清晰可辩，无干扰峰（图 3）。原始序列质量的评价使用软件 Sequencing Analysis 5.3.1（Applied Biosystems），即在 20 bp 的窗口内，平均 QV 值设定为≥20 时，修剪后的序列应长于原始序列读长的 50%。序列拼接时要求正反引物重叠序列区长于 100 bp，并且不确定的碱基数小于 1%（Li 等，2011b）。

对于比对好的序列，需要检查该序列中是否存在测序仪读取或编辑时造成的错误。对于编码基因（*rbc*L 和 *mat*K）来说，检查矩阵中是否有非 3 倍整数或大片段的插入/缺失，翻译的氨基酸是否有提前终止密码子等，排除假基因及可能的污染序列。对于非编码片段（ITS 和 *trn*H-*psb*A），由于种间可能存在较大的碱基或长度差异，矩阵排列可能相对困难，检查时应更仔细。首先对同种的不同个体进行排列，检查是否存在私有的变异位点（只发生在单个个体中），回去检查原始峰图确认该私有变异位点是否真实存在；然后再对近缘种进行排列，以相同的方法进行检查，这样可以最大限度地避免仪器错读或编辑时发生的错误。

10 DNA 条形码数据分析

DNA 条形码数据的分析使用最多的方法有 BLAST 法，距离法（distance）和建树法（Tree-building）。其中 BLAST 法得到的物种鉴定率最高，尤其在使用组合片段进行物种鉴定时，显著高于其它方法得到的物种鉴别率，而距离法与建树法得到的物种鉴别率相差不大（Li 等，2011a）。我们推荐使用距离法和建树法用于 DNA 条形码数据的分析。经序列比对和人工校对后，通过 MEGA 或 PAUP 等软件可计算种内和种间的 Kimura-2-parameter distance（K2P）遗传距离。根据该距离模型构建 NJ 树（Neighbour-joining tree）。

10.1 遗传距离计算

种间距离通常采用 pairwise uncorrected *p*-distance（Newmaster 等，2008）或 Kimura-2-parameter distance（K2P）（Meyer 和 Paulay，2005；Lahaye 等，2008）模型计算。K2P 距离是生物条形码联盟（CBOL）推荐使用的距离计算模型（barcoding.si.edu/）。种内遗传距离通常采用 3 种参数表示（Meyer 和 Paulay，2005；Lahaye 等，2008）：K2P 距离，平均 θ 值和平均溯祖度（average coalescent

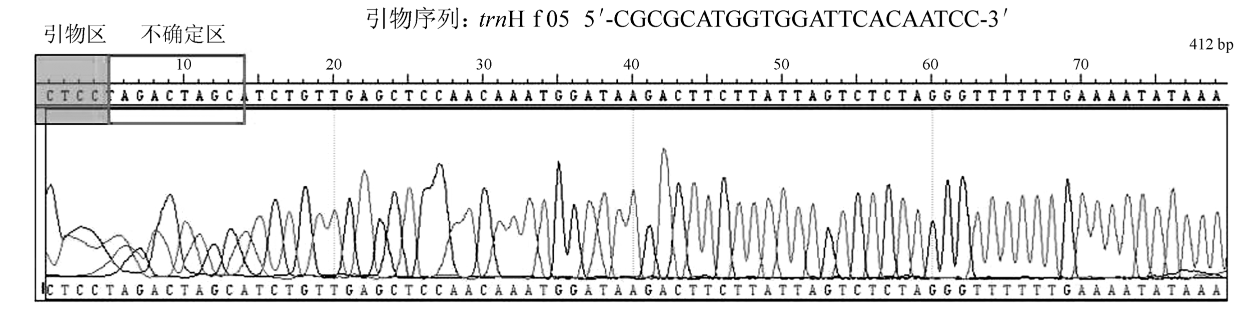


图 3 *trn*H-*psb*A 片段引物 *trn*H 的原始峰图。灰色标记区示引物区，白色标记框内示碱基不确定区

Fig. 3 The chromatogram of primer *trn*H for *trn*H-*psb*A region

The grey highlight region shows primer region, and white rectangle shows ambiguous region

depth)。其中平均 θ 值是指每个物种内不同个体间的平均 K2P 距离，目的是消除不同物种因采样个体数不均引起的偏差；平均溯祖度是指物种内所有个体间最大的 K2P 距离，用以反映种内最大变异范围。K2P 距离可以通过 MEGA 或 PAUP 计算，在此基础上计算其余两个参数。

10.2　系统树构建

DNA 条形码分析中通常采用标准的系统树构建方法（如 NJ、UPGMA、ML、MP、Bayes）。建树的目的并不是利用条形码重建系统发育树，而是为了检验每个物种的单系性，即同一物种的不同个体能否紧密聚类到一起。不同的建树方法可能得到不同的物种分辨率，但不同方法得到的结果差别并不大。此外，不同方法的运算时间差别很大，而且适用的条件不同，在使用时应根据需要进行选择。目前使用最多的建树方法是 NJ 法。

11　植物 DNA 条形码研究展望

DNA 条形码概念自 Herbert 等（2003）首次提出后，在近十年时间内得到了快速发展，已广泛应用于物种的分类和鉴定、发现新种和隐型种、生态监测及生物多样性编目等研究领域，表现出强大的生命力。虽然我国开展 DNA 条形码的研究起步较晚，特别是植物 DNA 条形码的研究到 2008 年才真正开始，但发展十分迅猛，发表了系列的相关文章，并参与了国际植物核心 DNA 条形码的评估与推荐，相信不久的将来我国在国际生命条形码研究中会发挥更大的作用。

中国是全球生物多样性极为丰富的国家之一，拥有约 10% 的全球植物物种，在全球生物多样性研究中占有重要的地位（Yang 等，2005）。2010 年 10 月日本名古屋联合国《生物多样性公约》第 10 次缔约方大会更新版《全球植物保护战略》（Global Strategy for Plant Conservation：GSPC）明确提出，到 2020 年完成世界在线植物志（World Flora Online）的战略目标，这既是机遇也是挑战。而 iFlora 将现代植物学、DNA 测序技术和信息技术的结合，通过系列的关键技术的集成和攻关，构建便捷、准确识别植物和掌握相关数字化信息的新一代植物志（李德铢等，2012），是在这一战略目标基础上的创新和发展，将表现出更大的生命力。通过相关关键技术的集成将极大促进植物分类学、分子系统学、演化生物学、生态学、保护生物学、生物地理学等学科的发展。

致谢　感谢英国爱丁堡皇家植物园 Pete Hollingsworth 教授和 Alan Forrest 博士提供的相关信息和有益讨论。

〔参　考　文　献〕

APG, 2009. An update of the Angiosperm Phylogeny Group classification for the orders and families of flowering plants: APG III [J]. *Botanical Journal of the Linnean Society*, **161**: 105—121

Baum DA, Small RL, Wendel JF, 1998. Biogeography and floral evolution of baobabs (*Adansonia*, Bombacaceae) as inferred from multiple data sets [J]. *Systematic Biology*, **47**: 181—207

Bergsten J, Bilton DT, Fujisawa T *et al.*, 2012. The effect of geographical scale of sampling on DNA barcoding [J]. *Systematic Biology*, **61**: 851—869

Bridson D, Forma L, 1992. *The Herbarium Handbook*, Revised edition [M]. London: Royal Botanic Gardens, Kew

CBOL Plant Working Group, 2009. A DNA barcode for land plants [J]. *Proceedings of the National Academy of Sciences of USA*, **106**: 12794—12797

Che J, Huang DW, Li DZ *et al.*, 2010. DNA barcoding and the International Barcode of Life Project in China [J]. *Bulletin of the Chinese Academy of Sciences*, **24**: 257—260

Chiou SJ, Yen JH, Fang CL *et al.*, 2007. Authentication of medicinal herbs using PCR-amplified ITS2 with specific primers [J]. *Planta Medica*, **73**: 1421—1426

Christenhusz MJM, Gardner MF, Mill RR *et al.*, 2011. A new classification and linear sequence of extant gymnosperms [J]. *Phytotaxa*, **19**: 55—70

Cuénoud P, Savolainen V, Chatrou LW *et al.*, 2002. Molecular phylogenetics of Caryophyllales based on nuclear 18S rDNA and plastid *rbcL*, *atpB*, and *matK* DNA sequences [J]. *American Journal of Botany*, **89**: 132—144

Doyle JJ, Doyle JL, 1987. A rapid DNA isolation procedure for small quantities of fresh leaf tissue [J]. *Phytochemical Bulletin*, **19**: 11—15

Dunning LT, Savolainen V, 2010. Broad-scale amplification of *mat*K for DNA barcoding plants, a technical note [J]. *Botanical Journal of the Linnean Society*, **164**: 1—9

Fay MF, Swensen SM, Chase MW, 1997. Taxonomic affinities of *Medusagyne oppositifolia* (Medusagynaceae) [J]. *Kew Bulletin*, **52**: 111—120

Fazekas AJ, Burgess KS, Kesanakurti PR *et al.*, 2008. Multiple multilocus DNA barcodes from the plastid genome discriminate plant

species equally well [J]. *PLoS One*, **3**: e2802

Fazekas AJ, Steeves R, Newmaster SG, 2010. Improving sequencing quality from PCR products containing long mononucleotide repeats [J]. *Biotechniques*, **48**: 277—285

Ford CS, Ayres KL, Toomey N *et al.*, 2009. Selection of candidate coding DNA barcoding regions for use on land plants [J]. *Botanical Journal of the Linnean Society*, **159**: 1—11

Gernandt DS, Liston A, 1999. Internal transcribed spacer region evolution in *Larix* and *Pseudotsuga* (Pinaceae) [J]. *American Journal of Botany*, **86**: 711—723

Hebert PDN, Cywinska A, Ball SL *et al.*, 2003. Biological identifications through DNA barcodes [J]. *Proceedings of the Royal Society of London Series B-Biological Sciences*, **270**: 313—321

Hebert PDN, Penton EH, Burns JM *et al.*, 2004. Ten species in one: DNA barcoding reveals cryptic species in the neotropical skipper butterfly *Astraptes fulgerator* [J]. *Proceedings of the National Academy of Sciences of USA*, **101**: 14812—14817

Hollingsworth PM, Graham SW, Little DP, 2011. Choosing and using a plant DNA barcode [J]. *PLoS One*, **6**: e19254

Kress WJ, Erickson DL, 2007. A two-locus global DNA barcode for land plants: the coding *rbc*L gene complements the non-coding *trn*H-*psb*A spacer region [J]. *PLoS One*, **2**: e508

Lahaye R, Van Der Bank M, Bogarin D *et al.*, 2008. DNA barcoding the floras of biodiversity hotspots [J]. *Proceedings of the National Academy of Sciences of USA*, **105**: 2923—2928

Li DZ, Gao LM, Li HT *et al.*, 2011a. Comparative analysis of a large dataset indicates that internal transcribed spacer (ITS) should be incorporated into the core barcode for seed plants [J]. *Proceedings of the National Academy of Sciences of USA*, **108**: 19641—19646

Li DZ, Liu JQ, Chen ZD *et al.*, 2011b. Plant DNA barcoding in China [J]. *Journal of Systematics and Evolution*, **49**: 165—168

Li Y, Gao LM, Poudel RC *et al.*, 2011c. High universality of *mat*K primers for barcoding gymnosperms [J]. *Journal of Systematics and Evolution*, **49**: 169—175

Li DZ (李德铢), Wang YH (王雨华), Yi TS (伊廷双) *et al.*, 2012. The next-generation Flora: iFlora [J]. *Plant Diversity and Resources* (植物分类与资源学报), **34** (6): 525—531

Liston A, Robinson WA, Oliphant JM *et al.*, 1996. Length variation in the nuclear ribosomal DNA internal transcribed spacer region of non-flowering seed plants [J]. *Systematic Botany*, **21**: 109—120

Liu J (刘杰), Gao LM (高连明), 2009. Comparative analysis of three different methods of total DNA extraction used in *Taxus* [J]. *Guihaia* (广西植物), **31**: 244—249

Liu J, Möller M, Gao LM *et al.*, 2011. DNA barcoding for the discrimination of Eurasian yews (*Taxus* L., Taxaceae) and the discovery of cryptic species [J]. *Molecular Ecology Resources*, **11**: 89—100

Liu J, Provan J, Gao LM *et al.*, 2012. Sampling strategy and potential utility of indels for DNA barcoding of closely related plant species: a case study in *Taxus* [J]. *International Journal of Molecular Sciences*, **13**: 8740—8751

Matz MV, Nielsen R, 2005. A likelihood ratio test for species membership based on DNA sequence data [J]. *Philosophical Transactions of the Royal Society B: Biological Sciences*, **360**: 1969—1974

Meyer CP, Paulay G, 2005. DNA barcoding: Error rates based on comprehensive sampling [J]. *PLoS Biology*, **3**: 2229—2238

Newmaster SG, Fazekas AJ, Steeves RAD *et al.*, 2008. Testing candidate plant barcode regions in the Myristicaceae [J]. *Molecular Ecology Resources*, **8**: 480—490

Ratnasingham S, Hebert PDN, 2007. BOLD: The Barcode of Life Data System (www. barcodinglife. org) [J]. *Molecular Ecology Notes*, **7**: 355—364

Sang T, Crawford DJ, Stuessy TF, 1997. Chloroplast DNA phylogeny, reticulate evolution and biogeography of *Paeonia* (Paeoniaceae) [J]. *American Journal of Botany*, **84**: 1120—1136

Stanford AM, Harden R, Parks CR, 2000. Phylogeny and biogeography of *Juglans* (Juglandaceae) based on *mat*K and ITS sequence data [J]. *American Journal of Botany*, **87**: 872—882

Tate JA, Simpson BB, 2003. Paraphyly of *Tarasa* (Malvaceae) and diverse origins of the polyploid species [J]. *Systematic Botany* **28**: 723—737

Valentini A, Pompanon F, Taberlet P, 2009. DNA barcoding for ecologists [J]. *Trends in Ecology and Evolution*, **24**: 110—117

White TJ, Bruns T, Lee S *et al.*, 1990. Amplification and direct sequencing of fungal ribosomal RNA genes for phylogenetics [A]. In: Innis MA, Gelfand DH, Sninsky JJ *et al.* (eds.), *PCR Protocols: A Guide to Methods and Applications* [M]. New York: Academic Press, Inc., 315—322

Yang QE, Zhu GH, Hong DY *et al.*, 2005. World's largest flora completed [J]. *Science*, **309**: 2163

Yao H, Song JY, Liu C *et al.*, 2010. Use of ITS2 region as the universal DNA barcode for plants and animals [J]. *PLoS One*, **5**: 370—375

Yu J, Xue JH, Zhou SL, 2011. New universal *mat*K primers for DNA barcoding angiosperms [J]. *Journal of Systematics and Evolution*, **49**: 176—181

Zhang AB, He LJ, Crozier RH *et al.*, 2010. Estimating sample sizes for DNA barcoding [J]. *Molecular Phylogenetic and Evolution*, **54**: 1035—1039